International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Liver Cirrhosis Stage Classification Using Ensemble Machine Learning Techniques

Jenisha Blessy R[1], Ms.Nishan A H[2]

[1]Student / Information Technology, Francis Xavier Engineering College, Tirunelveli, India

jenishablessyr.ug.21.it@francisxavier.ac.in

[2]Assistant Professor / Information Technology, Francis Xavier Engineering College, Tirunelveli, India

nishan@francisxavier.ac.in

Abstract :

Liver cirrhosis is a chronic and dangerous condition that, if not detected early, can result in serious and often fatal complications. Early detection allows for improved treatment strategies and intervention plans to slow the progression of cirrhosis and enhance patient outcomes. However, most patients cannot afford early detection due to the high costs, invasive procedures, and lengthy time requirements of traditional detection methods like biopsies and radiological examinations. To solve this problem, we have created a highly advanced system on the basis of machine learning algorithms to forecast the development of liver cirrhosis. Rather than relying on expensive and time-consuming tests, our model employs key indicators of health such as bilirubin, albumin, platelet count, and liver enzyme activity to determine the degree of cirrhosis. Our model is derived from a Voting Classifier that utilizes the strengths of CatBoost, XGBoost, and LightGBM to make predictions. On the basis of real-world patient data, the system is able to predict the stages correctly as No Cirrhosis, Early Cirrhosis, Moderate Cirrhosis, and Severe Cirrhosis. In order to make the technology user-friendly, we designed a web application using Flask that allows health professionals to input patient information and get real-time predictions. The web interface is convenient and simple to use and provides results for classification along with personalized medical recommendations. The user can also be given a detailed PDF report with the diagnosis, input medical parameters, and recommendations for future action in the form of further medical evaluation.In order to ensure the system works as expected in real life, the model is comprehensively tested against measures like precision, recall, accuracy, and F1-score. The accuracy of misclassification is checked with a confusion matrix, which helps improve the model's performance.ease diagnosis, and improves patient care through early treatment. This system's machine learning and user-friendly interface make it easier, faster, and more accurate for a wider range of users to identify liver cirrhosis. By enabling early interventions, this tool improves patient care, expedites physician decision-making, and facilitates early detection. It can be applied to preventive care, telemedicine, and hospitals.

Keywords - **Liver Cirrhosis Classification, Machine Learning, Voting Classifier, CatBoost, XGBoost, LightGBM, Predictive Analytics, Feature Selection, Disease Progression Analysis, Preventive Healthcare.**

I. INTRODUCTION

Liver cirrhosis is a serious and progressive illness that impacts millions globally. It develops when chronic liver damage results in scarring (fibrosis), which hinders the liver's ability to function effectively. Early detection is vital for successful treatment and improved patient outcomes; however, conventional diagnostic techniques like liver biopsies and imaging scans can be invasive, expensive, and time-consuming. Consequently, a lot of cases remain undetected until serious issues arise, making it tough to intervene early. This study is focused on developing a machine learning tool to classify the stages of liver cirrhosis. By analyzing key health indicators like bilirubin levels, albumin, platelet count, and liver enzyme levels, the system identifies patterns and accurately sorts cirrhosis into four categories: no cirrhosis, early cirrhosis, moderate cirrhosis, and severe cirrhosis. This model uses a voting classifier that combines the benefits of CatBoost, XGBoost, and LightGBM to enhance both the accuracy and reliability of predictions. To make it more user-friendly, the system is built into a flask-based web application, which allows healthcare professionals to enter patient details and receive quick predictions. The interface shows clear classification outcomes, confidence scores, and customized medical advice. Furthermore, a feature for generating PDF reports allows users to download and keep patient records for further medical assessment and treatment planning.By utilizing predictive analytics and real-time data processing, this project facilitates early detection, aids in making informed treatment choices, and improves the management of liver diseases. Its potential applications include clinical environments, telemedicine, and preventive healthcare, offering an innovative, data-driven strategy to enhance patient outcomes and alleviate the impact of liver cirrhosis globally.

II. LITERATURE REVIEW

[1] A. E. Topcu, E. Elbasi, and Y. I. Alzoubi, 2024. "Machine Learning-Based Analysis and Prediction of Liver Cirrhosis". This paper presents a machine learning-driven approach to analyzing and predicting liver cirrhosis, evaluating multiple models to enhance diagnostic accuracy.

[2] J. Allenki and H. Kumar. Soni, 2024. "Analysis of Chronic Liver Disease Detection by Using Machine Learning Techniques". This study explores various machine learning algorithms for chronic liver disease detection, focusing on feature selection and classification techniques.

[3] I. Hanif and M. M. Khan, 2022. "Liver Cirrhosis Prediction Using Machine Learning Approaches". This research compares different machine learning models to predict liver cirrhosis stages, emphasizing model performance metrics.

[4] R. Manjunath, A. Ghanshala, and K. Kwadiki, 2023. "Deep Learning Algorithm Performance Evaluation in Detection and Classification of Liver Disease Using CT Images". The paper discusses deep learning techniques for liver disease classification using CT images, providing a comparative analysis of different architectures.

[5] N. Nithyashri, H. Goel, and M. S. Hada, 2024. "Intelligent Classification of Liver Diseases using Ensemble Machine Learning Techniques". This paper explores the use of ensemble machine learning techniques for the classification of liver diseases, improving diagnostic accuracy through model integration.

[6] M. Suárez, S. Gil-Rojas, P. Martínez-Blanco, A. M. Torres, A. Ramón, P. Blasco-Segura, M. Torralba, and J. Mateo, 2024. "Machine Learning-Based Assessment of Survival and Risk Factors in Non-Alcoholic Fatty Liver Disease-Related Hepatocellular Carcinoma". The paper introduces a predictive model for assessing patient survival and risk factors associated with liver cancer.

[7] N. Han, J. He, L. Shi, M. Zhang, J. Zheng and Y. FanN, 2022. "Identification of Biomarkers in Nonalcoholic Fatty Liver Disease: A Machine Learning Method and Experimental Study". This study employs machine learning to identify key

biomarkers associated with nonalcoholic fatty liver disease.

[8] T. M. Ghazal, A. U. Rehman, M. Saleem, M. Ahmad, S. Ahmad and F. Mehmood, 2022. "Intelligent Model to Predict Early Liver Disease Using Machine Learning Technique". The paper presents an AI-based predictive model to detect early-stage liver diseases, emphasizing accuracy improvements.

[9] R. Amin, R. Yasmin, S. Ruhi, M. H. Rahman and M. S. Reza, 2023. "Prediction of Chronic Liver Disease Patients Using Integrated Projection-Based Statistical Feature Extraction with Machine Learning Algorithms". This research proposes a novel feature extraction approach combined with machine learning for chronic liver disease prediction.

[10] K. Gupta, N. Jiwani, N. Afreen and D. Divyarani, 2022. "Liver Disease Prediction Using Machine Learning Classification Techniques". This study evaluates different classification algorithms for liver disease prediction, emphasizing ensemble learning techniques.

III. METHODOLOGY

A. Understanding the Problem

Liver cirrhosis is a serious condition that slowly damages liver function, so catching it early is crucial for better treatment success. This project plans to create an online tool that takes users' health data and uses machine learning to predict how cirrhosis will progress. The aim is to help doctors make smart choices while providing an easy and accessible option for patients. By connecting healthcare with technology, this system improves early diagnosis and treatment.

B. Data Collection and Preprocessing

The dataset used for this project includes real medical records of patients diagnosed with liver cirrhosis, containing vital parameters like Bilirubin levels, Albumin, and Liver Enzymes. To achieve accurate data, we conduct a thorough cleaning process. This involves correcting errors, managing missing values, and standardizing formats. We focus on selecting crucial features that greatly impact cirrhosis predictions to boost the model's performance. Normalization techniques are employed to ensure consistency among different medical readings. These actions help produce a high-quality dataset, enabling the model to learn from the best available information.

C. Model Selection and Training

To find the most reliable approach, different machine learning algorithms like Random Forest, SVM, and Neural Networks are tested on the dataset.We separate the dataset into training and testing sections to accurately evaluate the model's performance. We split the dataset into training and test sets to ensure the model learns effectively and generates accurate predictions. To evaluate the effectiveness of the model, we employ important metrics like accuracy and recall. This helps us choose the most reliable model.

D. Building the Web Application

The web application is intended to be simple and easy for healthcare professionals to access. It uses Flask on the backend to manage user data and run the trained model for making predictions. On the frontend, HTML, CSS, and JavaScript combine to create a user-friendly form for entering medical information. After they submit the form, the system promptly evaluates the information and yields a prediction. The aim is to provide AI-driven health insights to all, regardless of their technical expertise.

E. Model Deployment and Integration

After training and testing the model, we integrate it into the Flask backend so it can make real-time predictions. We test each feature thoroughly to confirm that everything functions properly, from validating user inputs to delivering correct results. This process upgrades our project from a simple

prototype to a fully working tool that anyone can use on any device without needing to install anything.

F. Evaluation and Optimization

We keep the system up to date for better accuracy. Incorporating functionalities such as Power BI visualization will help in the explanation of the results. Regular updates keep the system up to date and easy to use for medical purposes.

IV. EXISTING SYSTEM

Currently, liver cirrhosis is primarily diagnosed by medical tests, physical examination, and a physician's examination. Blood tests monitor the functioning of the liver, while imaging techniques such as ultrasounds and MRI's provide a better idea of the condition of the liver. In certain cases, a liver biopsy is required for a more precise diagnosis. While these approaches are effective, they are time-consuming and expensive, and they require specialized medical equipment. A major drawback is that liver cirrhosis develops silently, and only later do symptoms appear. This delay makes the treatment less effective. Also, in remote areas, advanced diagnostic tools and healthcare workers are often hard to find. Some research groups and hospitals are looking into AI-based diagnostic technology, but it's still new and not commonly used yet. Many existing AI systems require complex installations and need to interface with hospital databases, so these are difficult to implement in general. This illustrates the requirement for a simpler AI-based system that can assist physicians and patients in detecting liver cirrhosis earlier, resulting in more effective treatment.

V. PROPOSED SYSTEM

A. Liver Cirrhosis Stage Classification Based on Data Analysis

The system gathers important health data like bilirubin levels, albumin, platelet count, and liver enzymes to assess liver cirrhosis severity. It categorizes cases into No Cirrhosis, Early Cirrhosis, Moderate Cirrhosis, and Severe Cirrhosis. Using machine learning, the system detects patterns in patient data to improve accuracy. The intention is to provide a trustworthy classification that helps in identifying diseases early and planning treatments more effectively. By looking at structured medical information, the system ensures accurate stage identification and aids in monitoring the condition.

B. Feature Extraction and Classification Using Machine Learning

The system employs a Voting Classifier that combines CatBoost, XGBoost, and LightGBM for accurate classification. Each algorithm brings its own unique benefits, which helps the model to accurately detect different stages of cirrhosis. The dataset is prepared beforehand to improve learning and enhance prediction accuracy. By fine-tuning the model, it can better understand complex connections in medical data. Using multiple algorithms makes the classification process more dependable and efficient. This strategy supports medical professionals in making better decisions by providing consistent and verified results.

C. Real-Time Data Processing and Prediction

The system processes patient data in real-time, delivering instant predictions on cirrhosis stages. Before we start analyzing, we check that the input data is modified and normalized to ensure it's correct. After you submit it, the model quickly reviews the health indicators and provides a classification result in just a few seconds. It also includes confidence scores that help users determine how reliable the results are, allowing them to make informed decisions. This system simplifies diagnosis by reducing the lengthy wait times common in conventional methods. Achieving results more rapidly leads to swifter medical assistance and enhanced patient treatment.

D. Intuitive Interface for Effortless Use

The system has an easy-to-use interface that is perfect for both healthcare workers and everyday

users. It's simple for users to enter patient information and get quick, clear results. The interface shows predictions along with confidence scores and recommended next steps to help users understand better. It's easy to navigate the system, so users can find important features without needing tech skills. The design creates a smooth experience, making AI health insights accessible to all. The user-friendly layout lets people focus on their decisions rather than getting bogged down by tech issues.

E. Optimizing performance and fine-tuning

To ensure the model stays accurate, it is frequently reviewed and enhanced. By including new patient data, the system becomes more skilled at learning and predicting outcomes. Key metrics such as accuracy, precision, recall, and F1-score are monitored to evaluate its performance. Cross-validation methods are used to make sure the model performs well with various datasets. To avoid overfitting, careful adjustments are made, ensuring that predictions are trustworthy in real-life situations. The aim is to continuously enhance the model for reliable and accurate diagnoses.

F. Stage Prediction System

The system gathers health information from patients, including bilirubin levels, albumin, platelet counts, and liver enzymes, to forecast the progression of liver cirrhosis. It classifies cases into four stages: No Cirrhosis, Early Cirrhosis, Moderate Cirrhosis, and Severe Cirrhosis. Using advanced machine learning techniques, it achieves a high level of accuracy in detecting the early stages of the disease. The organized dataset helps in accurately identifying the different stages of cirrhosis. This predictive method supports healthcare workers in making prompt treatment choices. The goal of the system is to enhance monitoring of liver diseases and strategies for early intervention.

G. Report Download

Users have the option to download a comprehensive PDF report that details their liver cirrhosis stage classification. This report gives a summary of the predicted stage and includes confidence

scores to help users understand the information better. It also features important health parameters that the system has analyzed, ensuring clear results. This research can help healthcare providers look for new diagnostic and therapeutic approaches. The option to download makes sharing easier and provides for quick reference in future sessions. The system ensures that reports are generated quickly and securely, allowing customers to easily monitor their health data.

VI. ARCHITECTURE DIAGRAM

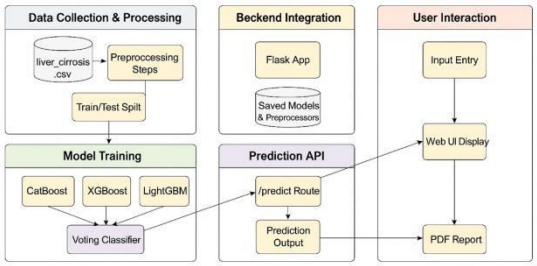

Fig 1 Architecture Diagram

A. Data collection and preparation

The system collects training and testing data, which consists of crucial medical records and important liver function metrics from patients. This data is organized by different stages of liver cirrhosis, allowing the model to learn properly and provide accurate predictions.

B. Data Cleaning and Enhancement

Before the training process begins, the dataset goes through several cleaning steps. This includes fixing missing information, adjusting the scale of features, and balancing the classes by creating synthetic data. These actions help make the model stronger and boost its accuracy in classification

C. Training the Model with a Voting Classifier

The organized data is processed by a Voting Classifier that integrates CatBoost, XGBoost, and LightGBM. This configuration enables the model to detect significant patterns associated with liver diseases through several training iterations, enhancing its accuracy.

D. Constructing and Assessing the Model

After training, the model undergoes testing using evaluation metrics like accuracy, precision, and recall. Analyzing its performance reveals areas that need enhancement and ensures dependable predictions.

E. Building and Evaluating the Model

Once the model is trained, it is evaluated using metrics like accuracy, precision, recall, and F1-score. This performance review helps ensure that the predictions are trustworthy and highlights areas that could be improved.

F. Integration and Launch with Flask

The trained model is incorporated into a web application using Flask, creating an easy-to-use interface for healthcare professionals to enter patient information and get predictions.The frontend is developed using HTML, CSS, and JavaScript.

G. User Interaction and Prediction

Users enter patient details via the web interface, and the model evaluates the data to determine the stage of liver cirrhosis. The system provides instant results, enabling early diagnosis and treatment planning.

VII. EXPERIMENTAL RESULTS

Users begin their journey on the home page, which has a friendly design and a noticeable "Get Started" button. By clicking this button, users are easily guided to the prediction page, where they can input important information about the patient.

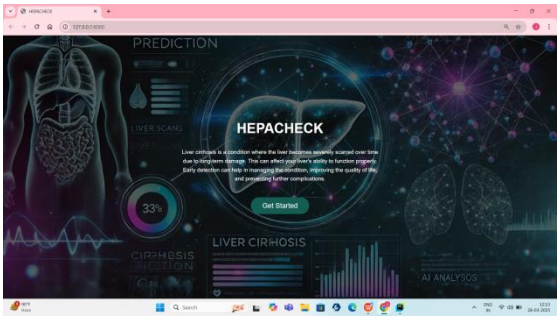

Fig 1 Home Page

On the prediction page, there are spaces to enter key details like Patient Name, Age, Gender, Total Bilirubin, Alkaline Phosphatase, Albumin, Prothrombin, Platelets, SGOT, Cholesterol, Triglycerides, Copper, Ascites, Hepatomegaly, Spiders, and Edema. The form is organized to make it easy for users to input and check their information before moving on.

After entering all the necessary details, users can hit the "Predict" button, which activates the machine learning model to analyze the data. The model quickly processes the information and gives an immediate prediction about the stage of liver cirrhosis.

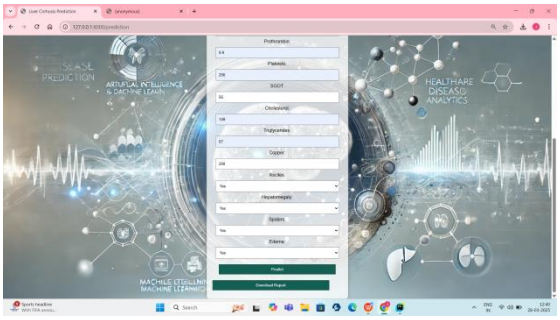

Fig 2 Prediction Page

Once the prediction is made, a "Download Report" button shows up, letting users create a detailed PDF report. The report is automatically named using the format`{patient_name}_Liver_Cirrhosis_Report.pdf`, making it personalized.

This report includes all the patient details, the predicted stage of liver cirrhosis, and important medical insights. This way, users get a clear and professional summary of their results, which they can share with doctors for further assessment.

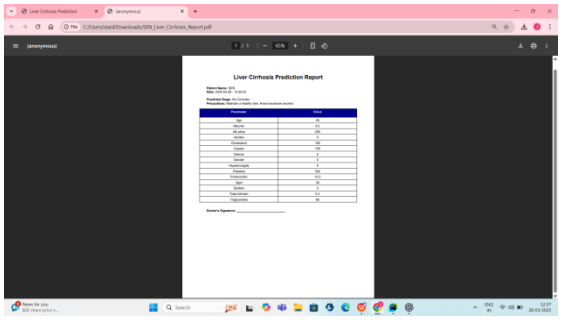

Fig 3 Downloaded PDF Report

This system is built to be simple, quick, and easy for everyone to use. The smooth transitions between pages, fast data processing, and automatic report generation improve the experience, making it both helpful and educational.

VIII. CONCLUSION

The project shows how well machine learning can classify the stages of liver cirrhosis. To protect patient information and ensure data safety, it features a secure login system that allows healthcare workers to evaluate liver health safely. A key benefit of this project is its ability to provide real-time prediction, users can enter patient information and get immediate feedback on how severe the cirrhosis is. By automating the classification, the system makes diagnosing easier and lessens the workload for healthcare professionals, leading to quicker and more informed decisions. Plus, the user-friendly web interface makes it simple for medical staff to understand and use the results without needing a lot of technical skills. This project also sets the stage for future improvements in liver disease detection by looking into ways to combine different types of data, use cloud technology, monitor with wearable devices, and support telemedicine. These upgrades could greatly enhance the system's accuracy, ease of use, and ability to grow, ultimately helping with early detection, preventive care, and better outcomes for patients.

XI. FUTURE SCOPE

Liver cirrhosis is divided into four stages: No Cirrhosis, Early Cirrhosis, Moderate Cirrhosis, and Severe Cirrhosis. However, with advancements in technology, we can enhance our understanding by including more health indicators such as genetic information, liver stiffness tests, and imaging scans. Expanding the variety of patient data we collect will help us find rare cases more effectively.

To make our predictions more accurate, we can merge lab test results with a patient's medical history, symptoms, and demographic information. This approach allows us to spot early warning signs that might not show up in regular medical tests. A thorough and personalized report would enable doctors to develop more effective treatment plans tailored to each individual.

A mobile app could elevate this system by providing doctors and healthcare providers with immediate access to patient reports, regardless of their location. This would help them make quick and informed decisions. A cloud-based system could further improve efficiency by securely handling large amounts of patient data in real time.

Linking this technology with telehealth platforms would be a significant advancement. Doctors could analyze patient data and offer medical advice remotely, which would lessen the need for in-person appointments. Connecting the system to electronic health records would help in easily tracking and monitoring patient histories, which is important for providing consistent care.

Since patient data is very private, security should be a major focus. Strong encryption and secure login methods, like using fingerprints or facial recognition, can help keep records safe from unauthorized access. Blockchain technology plays a vital role in keeping patient information safe and stopping unauthorized changes. It makes sure that records are secure, giving both doctors and patients peace of mind.

As we look to the future, combining this technology with wearable health devices could create exciting new opportunities. These devices can consistently monitor vital health indicators like heart rate and oxygen saturation, promptly alerting you if something appears unusual. By recognizing possible health issues early, they can prevent severe complications, shifting healthcare towards preventing problems before they occur.

X. REFERENCES

[1] A. E. Topcu, E. Elbasi and Y. I. Alzoubi, "Machine Learning-Based Analysis and Prediction of Liver Cirrhosis," 2024 47th International Conference on Telecommunications and Signal Processing (TSP), Prague, Czech Republic, 2024, pp. 191-194.

[2] J. Allenki and H. Kumar. Soni, "Analysis of chronic liver disease detection by using machine learning techniques", 2024 IEEE International Students' Conference on Electrical Electronics and Computer Science (SCEECS), pp. 1-8, 2024.

[3] I. Hanif and M. M. Khan, "Liver cirrhosis prediction using machine learning approaches", 13th Annual Ubiquitous Computing Electronics & Mobile Communication Conference (UEMCON), pp. 0028-0034, 2022.

[4] R. Manjunath, A. Ghanshala and K. Kwadiki, "Deep learning algorithm performance evaluation in detection and classification of liver disease using CT images", Multimedia Tools and Applications, vol. 83, no. 1, pp. 2773-2790, 2023.

[5] N. Nithyashri, H. Goel, and M. S. Hada, "Intelligent classification of liver diseases using ensemble machine learning techniques," 2nd Int. Conf. on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Aug. 2024.

[6] M. Suarez et al., "Machine learning-based assessment of survival and risk factors in non-alcoholic fatty liver disease-related hepatocellular carcinoma for optimized patient management", Cancers, vol. 16, no. 6, pp. 1114, 2024.

[7] N. Han, J. He, L. Shi, M. Zhang, J. Zheng and Y. Fan, "Identification of biomarkers in nonalcoholic fatty liver disease: A machine learning method and experimental study", Frontiers in Genetics, vol. 13, pp. 1020899, 2022.

[8] T. M. Ghazal, A. U. Rehman, M. Saleem, M. Ahmad, S. Ahmad and F. Mehmood, "Intelligent model to predict early liver disease using machine learning technique", 2022 International Conference on Business Analytics for Technology and Security (ICBATS), pp. 1-5, 2022.

[9] R. Amin, R. Yasmin, S. Ruhi, M. H. Rahman and M. S. Reza, "Prediction of chronic liver disease patients using integrated projection-based statistical feature extraction with machine learning algorithms", Informatics in Medicine Unlocked, vol. 36, pp. 101155, 2023.

[10] K. Gupta, N. Jiwani, N. Afreen and D. Divyarani, "Liver disease prediction using machine learning classification techniques", 11th International Conference on Communication Systems and Network Technologies (CSNT), pp. 221-226, 2022.

[11] N. Slivinski and Z. Sheikh, "Cirrhosis of the liver: Symptoms, stages, and treatment", [online] Available: https://www.webmd.com/digestive-disorders/understanding-cirrhosis-basic-information.

[12] R. A. Khan, Y. Luo and F.-X. Wu, "Machine learning based liver disease diagnosis: A systematic review", Neurocomputing, vol. 468, pp. 492-509, 2022.

[13] M. Noureddin et al., "Machine learning liver histology scores correlate with portal hypertension assessments in nonalcoholic steatohepatitis cirrhosis", Alimentary Pharmacology & Therapeutics, vol. 57, no. 4, pp. 409-417, 2023.

[14] J. Tian, R. Cui, H. Song, Y. Zhao and T. Zhou, "Prediction of acute kidney injury in patients with liver cirrhosis using machine learning models: Evidence from the MIMIC-III and MIMIC-IV", International Urology and Nephrology, vol. 56, no. 1, pp. 237-247, 2024.

[15] M. P. Behera, A. Sarangi, D. Mishra and S. K. Sarangi, "A hybrid machine learning algorithm for heart and liver disease prediction using modified particle swarm optimization with support vector machine", Procedia Computer Science, vol. 218, pp. 818-827, 2023.

[16] G. Chongo and J. Soldera, "Use of machine learning models for the prognostication of liver transplantation: A systematic review", World Journal of Transplantation, vol. 14, no. 1, 2024.

[17] E. Dritsas and M. Trigka, "Supervised machine learning models for liver disease risk prediction", Computers, vol. 12, no. 1, pp. 19, 2023.

[18] A. Q. Md, S. Kulkarni, C. J. Joshua, T. Vaichole, S. Mohan and C. Iwendi, "Enhanced preprocessing approach using ensemble machine learning algorithms for detecting liver disease", Biomedicines, vol. 11, no. 2, pp. 581, 2023.

[19] R. K. Sachdeva, P. Bathla, P. Rani, V. Solanki and R. Ahuja, "A systematic method for diagnosis of hepatitis disease using machine learning", Innovations in Systems and Software Engineering, vol. 19, no. 1, pp. 71-80, 2023.